

WÖRTERBÜCHER ALS SEMANTISCHE NETZE

Einleitung

Der Wortschatz einer Sprache ist ein schier undurchdringliches Dickicht an Bedeutungsnuancen, Synonym- und Antonymbeziehungen, Hierarchien, Metonymien, speziellen Verwendungsweisen und vielerlei „Fallstricken“ mehr. Muttersprachlerinnen und Muttersprachler bewegen sich die meiste Zeit anscheinend mühelos durch diesen Dschungel des eigenen Wortschatzes – gelegentliche Ausrutscher ausgenommen.

Lernen wir eine uns fremde Sprache, können wir uns im fremden Wortschatz-Dschungel manchmal noch an einzelnen Wörtern orientieren, die so klingen oder aussehen wie Wörter unserer Muttersprache. Doch auch hier ist Vorsicht geboten: Die US-amerikanische „billion“ ist tausendmal weniger als die deutsche „Billion“ – um nur eines von vielen Beispielen zu nennen. Doch die meiste Zeit sind wir eher hilflos und nehmen nur allzu oft den falschen Abzweig, wenn wir uns in einer für uns neuen Sprache verständlich machen wollen. Hier können uns Wörterbücher Orientierung bieten und uns wichtige Hinweise geben, wie wir uns in und mit der Sprache sicherer bewegen können. Ein Wörterbuch macht uns das System der uns fremden Sprache übersichtlicher, indem es gegebenenfalls gezielt auf Fallstricke hinweist oder Schlaglichter setzt auf uns bisher unbekannte Verwendungsweisen eines Worts, letzteres funktioniert natürlich auch in der eigenen Muttersprache.

UNSER WORTSCHATZ IST EIN DICHTES GEFLECHT VON BEZIEHUNGEN

Und noch eine Eigenschaft von Wörterbüchern ist hoch zu schätzen: Die Serialität, also das alphabetische Eins-nach-dem-anderen, das die Wörterbuchartikel ordnet, ist ein sinnvolles ausdrucksseitiges Arrangement. Inhaltsseitig aber stehen Wörter nach einer semantischen Ordnung zueinander in Beziehung. Die meisten Wörterbücher sind deshalb gespickt mit Verweisen, die uns oft quer durch das Wörterbuch schicken. Diese Querverweise können unterschiedlichste Funktionen erfül-

len. Sie können uns beispielsweise auf einen weiteren Wortartikel hinweisen, der ein Synonym behandelt. Oder es wird auf einen Artikel zu einer Mehrwortverbindung verwiesen, die unser augenblicklich betrachtetes Wort enthält. So könnte eine Engländerin, die gerade Deutsch lernt und sich über das deutsche Wort „Gras“ informiert, zum Beispiel entdecken, dass man auf Deutsch „ins Gras beißt“, wenn man auf Englisch „den Eimer tritt“ (*to kick the bucket*¹). Sind diese Verweise zwischen verschiedenen Einträgen im Wörterbuch für einen einzelnen Eintrag vielleicht noch recht überschaubar, so ergibt sich für den gesamten Wortschatz einer Sprache, oder besser: für den vom Wörterbuch beschriebenen Ausschnitt des Wortschatzes, ein dichtes Geflecht von Vernetzungen zwischen Wörtern.

WÖRTERBÜCHER KÖNNEN UN DURCHSICHTIGE ZUSAMMENHÄNGE AUFDECKEN

Das ist genau die Perspektive, die wir in diesem Artikel einnehmen wollen: das Wörterbuch als Netz, das über Ausschnitte des Wortschatzes gelegt wird und so bisher undurchsichtige Zusammenhänge offenlegt. Eine solche Betrachtungsweise mag zunächst nur als bloße Spielart traditioneller Perspektiven daherkommen. Wir wollen jedoch in diesem Beitrag ausführen, welchen Erkenntnisgewinn diese Netz-Metapher mit sich bringen kann. Dabei interessieren wir uns in letzter Konsequenz für jene Informationen und Schlussfolgerungen, die erst dadurch gewonnen werden können, dass wir Wörterbücher durch die Linse dieser Netz-Metapher betrachten.

Die Autoren sind wissenschaftliche Mitarbeiter der Abteilung Lexik am Institut für Deutsche Sprache in Mannheim.

Wir werden dabei wie folgt vorgehen: Wir werden zunächst das Wörterbuch vorstellen, anhand dessen wir unsere Perspektive aufzeigen werden. Dies ist das am IDS erarbeitete Diskurswörterbuch zum Schulddiskurs 1945-1955, das von Heidrun Kämper konzipiert und erstellt wurde (Kämper 2005, 2007). Im nächsten Schritt werden wir die Operationalisierung der Netz-Metapher erarbeiten und zeigen, wie dieses Bild anhand des mathematischen Konzepts eines Graphen operationalisiert werden kann. Der Hauptteil dieses Beitrags besteht aus der Anwendung der Graphentheorie auf das Diskurs-Wörterbuch. In diesem Abschnitt werden wir zeigen, wie das Netz aussieht, das vom Wörterbuch über einen ganz bestimmten Ausschnitt des deutschen Wortschatzes in einer bestimmten Zeitepoche gelegt wird. Wir werden weiterhin die konkreten Erkenntnisse, die aus dieser Abbildung hervorgehen, erarbeiten und interpretieren. Im letzten Abschnitt werden wir die gewonnenen Erkenntnisse in einem größeren Rahmen diskutieren und werden unter anderem zeigen, wie durch eine solche Integration von qualitativer und quantitativer Forschung neue Perspektiven auf Wörterbücher und den Wortschatz als Ganzes gewonnen werden können.

Diskurs-Wörterbuch zum Schulddiskurs 1945-1955

Zunächst ist zu klären: Was ist ein Diskurs? Ein Diskurs

- besteht aus Serien öffentlicher themenkohärenter, kommunikativer Akte: Diese thematisch-inhaltliche Bindung als Kohärenzfaktor ist diskursives Grundelement. Das thematische ist damit zugleich das Hauptkriterium zur Abgrenzung eines Diskurses;
- wird von Diskursbeteiligten realisiert: Sie konstituieren die Sprechergruppe, die die Beschaffenheit spezifischer, zeittypischer Diskurse bestimmt und ihnen Dichte und Dynamik gibt. Die Redeweisen der Diskursbeteiligten werden durch deren Zugehörigkeit zu einer Domäne bzw. Beteiligungsrolle geprägt;

- verdichtet sich in Leit- und Schlüsselwörtern, die untereinander in engen Bedeutungsbeziehungen stehen und die dem Diskurs seine lexikalisch-semantische Ordnung geben. Ihr Gebrauch ist gesellschaftlich und historisch gebunden. Solche aus dem Diskurszusammenhang rekonstruierte Semantik geht über die traditionelle lexikalische Semantik hinaus, denn: Es ist kontextbezogene Semantik – im Hinblick auf das Thema, die Beteiligten und ihre kommunikativen Absichten sowie hinsichtlich der Zeit / des Zeitpunkts ihrer Aktualisierung.

Diese Faktoren lassen erkennen, dass ein Diskurs ein gesellschaftliches (und damit sprachliches und historisch geprägtes) komplexes und semantisch geordnetes kohärentes Phänomen ist.

Diskurswörterbuch typologisch

Wo ist ein so konzipiertes Diskurswörterbuch in der Wörterbuchtypologie zu platzieren? Ein Diskurswörterbuch verzeichnet die lexikalisch repräsentierte Wirklichkeitssicht und -erfahrung nicht einer allgemeinen, sondern einer eingegrenzten spezifischen Gruppierung einer Sprachgesellschaft, der Diskursbeteiligten. Insofern hat ein Diskurswörterbuch einen niedrigeren Allgemeingültigkeitsgrad als ein allgemeinsprachliches Wörterbuch. Da es andererseits die Wirklichkeitssicht und -erfahrung nicht eines Individuums, sondern eines bestimmten Kollektivs wiedergibt, hat es andererseits einen höheren Allgemeingültigkeitsgrad als ein individualsprachliches, z. B. ein Autorenwörterbuch. Diskurslexikografie hat also ihren Platz zwischen dem allgemeinsprachlichen, weitgehend kontextfernen Standardwörterbuch und dem kontextnahen Spezialwörterbuch. Ein Diskurswörterbuch stellt dementsprechend die semantische Prägung lexikalischer Einheiten unter den Bedingungen des Diskurses dar, ohne allgemein-semantische unspezifi-

sche und ohne individualesemantische restriktive Aussagen über den jeweiligen Wortgebrauch zu machen. Insofern ist es keine Alternative, sondern eine Ergänzung eines allgemeinsprachlichen Wörterbuchs.

DISKURSWÖRTERBÜCHER REKONSTRUIEREN DIE SEMANTISCHE STRUKTUR EINES WORTBESTANDS AUF HOHER PRÄZISIONSEBENE

Denn indem ein Diskurswörterbuch über den Sprachgebrauch zu einer bestimmten Zeit durch bestimmte Beteiligte zu einem bestimmten Thema des öffentlichen Diskurses informiert, rekonstruiert es die semantische Struktur eines Wortbestands auf einer hohen Präzisionsebene.

Prinzipien der Diskurslexikologie und Diskurslexikografie

Lexikologisch bzw. lexikalisch-semantisch gesehen sind Diskurse Archive von Inhalts- und Ausdruckseinheiten, die durch ihre jeweils spezifischen historischen, gesellschaftlichen, politischen Bedingungen geprägt sind. Diese Inhalts- und Ausdruckseinheiten, also das Diskursvokabular, sind die diskursiven Verdichtungen und Ordnungseinheiten eines Diskurses.

Wir können also sagen: Die Struktur eines Diskurswortschatzes besteht aus der Summe des semantischen Potenzials der lexikalischen Diskurselemente und der Summe der regelmäßigen Beziehungen, die diese lexikalischen Diskurselemente untereinander eingehen.

DISKURSWÖRTERBÜCHER BESCHREIBEN DISKURS-KONSTITUIERENDES UND -REALISIERENDES VOKABULAR

Was ist unter diesen Voraussetzungen ein Diskurswörterbuch? Die Aufgabe eines Diskurswörterbuchs ist es, dasjenige Vokabular zu beschreiben, das diskursrelevant ist: das einen Diskurs konstituiert und ihn realisiert. Diese Beschreibung folgt spezifischen lexikografischen Prinzipien. Unter der Perspektive einer graphenorientierten Wortschatzdarstellung sind insbesondere die folgenden Prinzipien relevant:

Prinzip der Kontextsemantik

Diskursanalyse beschreibt und erklärt Funktionen von Diskursen in der Gesellschaft. ‚Kontextsemantik‘ meint das kontextabhängige, über einen zentralen Gegenstand definierte, vielfältig spezifizierte und aspektualisierte Bedeutungsgefüge, das einem Diskurs Kohärenz verleiht und das von den Diskursbeteiligten geschaffen wird. Ein Diskurswörterbuch unterscheidet sich daher von einem allgemeinsprachlichen Wörterbuch, weil es thematische, sprecherbezogene, zeitliche, funktionale und textliche Gebundenheiten des Wortschatzes darstellt.

Beteiligtenprinzip

Die Diskursbeteiligten sind die Vermittlungsinstanzen des Diskurses und konstituieren seine lexikalische Ordnung. Ihre sprachlichen Kodierungen des Diskursthemas sind geprägt von der Beteiligungsrolle, die sie im Diskurs haben, von ihrem Weltbild und ihrer Einstellung, von der Domäne, aus deren Perspektive sie sich am Diskurs beteiligen, also von fachlich-professionellen, religiösen, ethischen usw. Zugehörigkeiten. Die Feststellung der beteiligungsrollen- und domänenspezifischen Diskursbeteiligten und ihre adäquate Präsenz in einem ausgewählten Diskurskorpus ist insofern obligatorisch. Denn es sind die Beteiligten, die die Beschaffenheit von Diskursen bestimmen. Ein adäquates Korpus eines Diskurses ist ein solches, das auf der Grundlage historischer und gesellschaftlicher Kenntnisse ausgewählt ist.

Diskursrelevanzprinzip

Ein Diskurswörterbuch ist kein Schlagwörterbuch. Sein Lemmaansatz bezieht sich insofern nicht nur auf die meist substantivischen Haupt-Wörter eines Diskurses, sondern auf den gesamten diskursrelevanten Wortschatz. Diskursrelevanter Wortschatz kann z.B. in Umbruchzeiten auch aus Temporaladverbialia bestehen (wie *heute, jetzt, in diesen Tagen*), und Diskursrelevanz können auch Wortbildungseinheiten haben (*mit-, gegen-*).

Netzprinzip

Die lexikalischen Elemente eines Diskurses stehen zueinander in Bedeutungsbeziehungen. Insofern sind Diskurse semantische Netze, die diese Bedeutungsbeziehungen ihres lexikalischen Inventars konstituieren. Das Wesen eines Diskurswortschatzes besteht darin, dass er einen begrifflichen bzw. semantischen kohärenten Komplex bildet, dessen Einheiten durch unterschiedliche Bedeutungsrelationen aufeinander bezogen sind. Vorzustellen ist er als Netz, dessen Knoten, die Lemmata, in unterschiedlicher Intensität in unterschiedlicher Relation zu den lexikalischen Repräsentanten des Diskurses stehen. Das bedeutet, dass die den Diskurs konstituierenden Lemmata durch eine Verweisstruktur lexikalisch-semantische Bezüge herstellen. Diejenigen mit der dichtesten / ‚lebhaftesten‘ Verweisstruktur bilden semantische / funktionale Zentren. Das Netz des Diskursvokabulars gibt dem Diskurs also eine semantische Ordnungsstruktur. Diese im weitesten Sinn onomasiologische Ordnungsstruktur betrifft z. B. Synonymierelationen, Antonymierelationen, Hyperonymie- / Hyponymierelationen, Funktionsrelationen. Damit bildet das Diskurswörterbuch die Struktur eines Diskurses auf lexikalisch-semantischer Ebene mit einem Verweissystem ab. Das Verweissystem ist die Darstellung eines durch die lexikologisch-semantische Analyse eines Diskurses erlangten Ergebnisses.

Im Wörterbuch zum Schulddiskurs 1945-1955 und im Wörterbuch zum Protestdiskurs 1967/68 sind diese Prinzipien umgesetzt. Sie folgen dem Kontextprinzip, indem Texte aus den Jahren 1945 bis 1955 bzw. 1967/68 ausgewertet und deren lexikalisch-semantischen Befunde als Kodierungen unter den spezifischen Bedin-

gungen einer Nachkriegszeit bzw. einer hochpolitisierten Zeit wie die der späten 1960er Jahre analysiert wurden. Die Umsetzung des Beteiligtenprinzips bedeutet für die Nachkriegszeit, einen nach Opfern, Tätern und Nichttätern bzw. – für die späten 1960er Jahre – nach intellektueller und studentischer Linken zu unterscheidenden Diskurs zu rekonstruieren. Dem Relevanzprinzip wurde in Bezug auf die diskurslexikografische Darstellung u. a. der zeittypischen Temporaldeiktika bzw. von Wortbildungselementen entsprochen.

Die Umsetzung des Netzprinzips ist Gegenstand der anschließenden Darstellung, die sich auf das Wörterbuch zum Schulddiskurs bezieht.

Mathematische Operationalisierung der Netz-Metapher: Graphentheorie

Der Grundgedanke der Graphentheorie ist eigentlich recht einfach (für eine detaillierte Einführung in die Graphentheorie s. Turau 2009): Ein Graph besteht nur aus zwei Arten von „Bauteilen“: Knoten und Kanten. Knoten sind dabei irgendwelche diskreten Entitäten, d. h. Dinge, die sinnvoll als verschieden oder abgrenzbar bezeichnet werden können. Kanten sind die Verbindungen zwischen diesen Entitäten. Das klingt alles sehr allgemein, und das ist auch gewollt, denn Graphen können auf eine Vielzahl von Systemen angewendet werden. Stellen wir uns beispielsweise ein sehr einfaches Schienennetz vor. Hier ist jede Stadt, die einen Bahnhof hat, ein Knoten im Graphen. Eine Schienenverbindung zwischen zwei Städten ist eine Kante. Selbstverständlich muss nicht jede Stadt zu jeder anderen Stadt eine Schienenverbindung haben. Abbildung 1 könnte ein solches Schienensystem darstellen. Einige Dinge fallen hier direkt auf: Stadt D scheint in diesem System sehr wichtig zu sein, fast alle Städte (bis auf

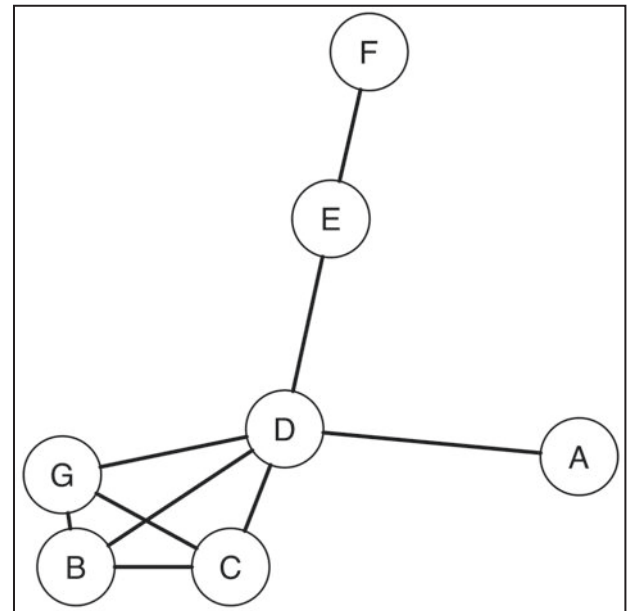


Abb. 1: Ein Graph mit 7 Knoten (Städten) und 9 Kanten (Schienenverbindungen)

Stadt F) haben eine direkte Verbindung zu Stadt D. Die Graphentheorie stellt für diese Beobachtung eine bestimmte Maßzahl zur Verfügung: Man sagt, dass der (Vernetzungs-)Grad des Knotens D der höchste im Graphen ist. Der Grad eines Knotens wird ganz einfach ermittelt: Man zählt einfach die Anzahl der Kanten, die an einem Knoten anliegen. Stadt D in Abbildung 1 hat einen Grad von 5, weil 5 Kanten (zu A, B, C, E und G) direkt an D anliegen. Die Stadt E hat hingegen einen Grad von 2, Stadt B einen Grad von 3 und so weiter.

Es fällt außerdem auf, dass die Städte B, C, D und G untereinander sehr gut verbunden sind, denn jede Stadt in dieser Gruppe hat eine direkte Verbindung zu jeder anderen Stadt in dieser Gruppe. Die Graphentheorie stellt auch hierfür ein Konzept bereit: Die Städte B, C, D und G bilden eine sogenannte Clique. Cliques sind Mengen von Knoten, die untereinander perfekt vernetzt sind, d. h., jeder Knoten ist mit jedem Knoten direkt verbunden. Man sieht sehr schnell, dass Stadt A nicht Teil dieser Clique ist, denn Stadt A ist nur mit Stadt D direkt verbunden, aber mit keiner der Städte B, C oder G.

GRAPHEN SIND VIELSEITIGE BESCHREIBUNGSTRUMENTE

An dieser Stelle ist uns wichtig darauf hinzuweisen, dass das imaginäre Beispiel von sieben Städten und den Schienenverbindungen zwischen ihnen beliebig gewählt ist. Graphen können immer dann als Beschreibungsinstrument herangezogen werden, wenn die

Verbindungsstruktur zwischen abgrenzbaren Einheiten in den Fokus rücken soll. Graphen finden sich in vielen Bereichen unseres Lebens. Die Knoten können dabei beispielsweise auch Funktionsstellen in einer Organisation sein und die Verbindungen könnten Weisungsbefugnis symbolisieren. In diesem Fall wäre der Graph eine spezielle Art eines Organigramms. Elektrische Schaltpläne oder die Dateisystemhierarchie auf dem PC sind Graphen. Auch ein Stammbaum ist ein Graph: Ein Knoten steht dann für einen Menschen und die Verbindungen zwischen Knoten symbolisieren verwandtschaftliche Beziehungen. Graphentheoretische Betrachtungsweisen sind auch in der Linguistik nichts Neues: So können auch Bäume, die die syntaktische Struktur eines Satzes beschreiben, als Graphen verstanden werden. Die verschiedenen Bestandteile eines Satzes (Nominalphrasen, Verbalphrasen, Präpositionalphrasen usw.) und die Wörter bilden dabei die Knoten des Graphen. Bäume sind lediglich ganz bestimmte Formen eines Graphen, die monohierarchische Beziehungen darstellen.

In diesem Beitrag wollen wir die Graphentheorie auf ein Feld der Linguistik anwenden, das bisher noch nicht auf diese Art betrachtet wurde – die Lexikografie. Die Übertragung ist dabei denkbar einfach: Einzelne Wörterbuchartikel sind die Knoten des Graphs, denn sie sind untereinander gut abgrenzbar. Wann immer in einem Artikel auf einen anderen verwiesen wird, gilt das als Verbindung zwischen den beiden Artikeln. Im vorliegenden Artikel werden wir dabei nicht auf die Verweisrichtung eingehen. Die Graphentheorie erlaubt es aber grundsätzlich, die Richtung einer Verbindung zu beachten. Es wäre dann ein Unterschied, ob von Artikel A zu Artikel B verwiesen wird oder andersherum.

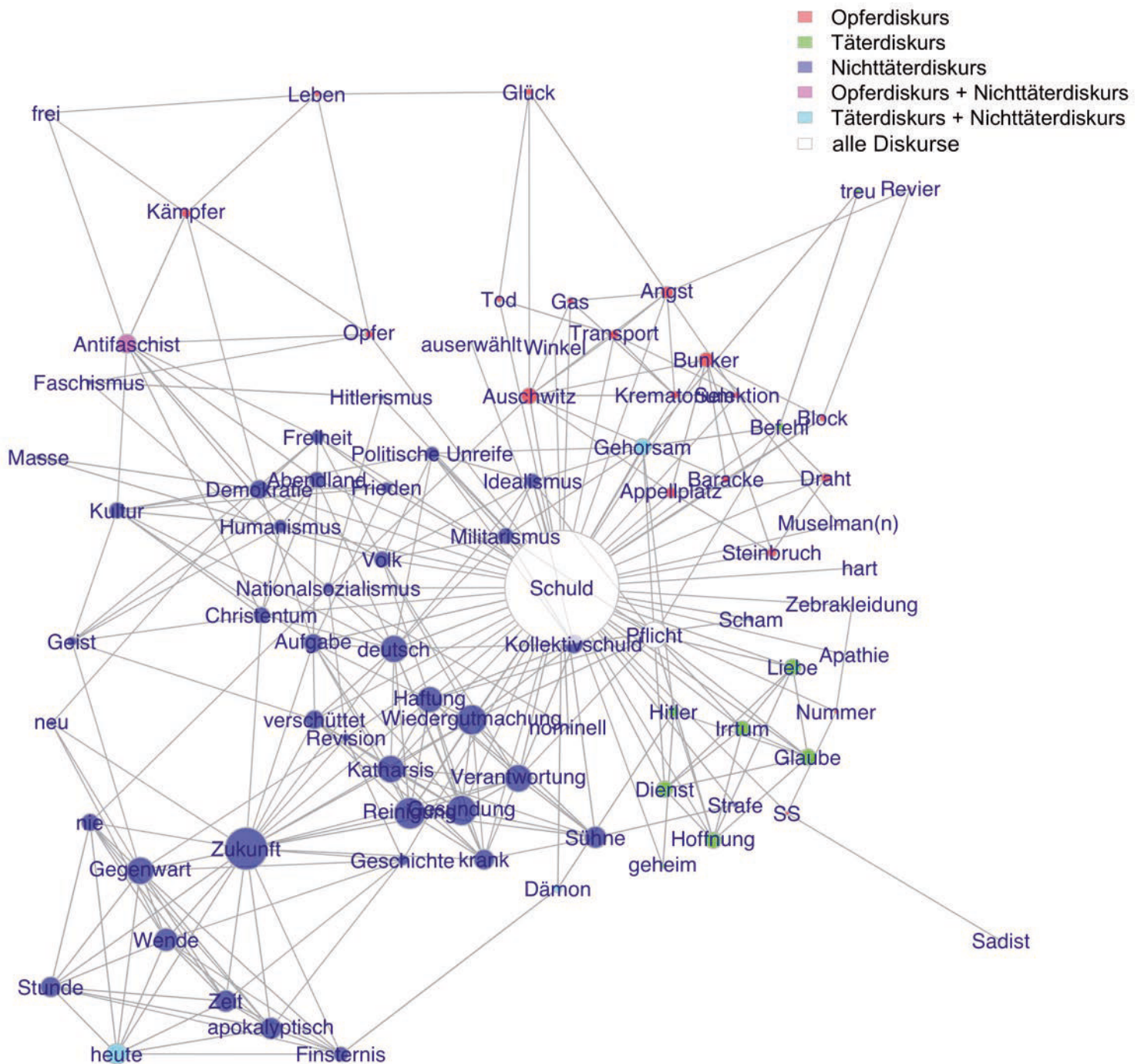


Abb. 2: Der Wörterbuchgraph zum Diskurswörterbuch Schuldiddiskurs 1945-1955. Jeder Knoten symbolisiert einen Wörterbuchartikel. Jede Kante symbolisiert einen Verweis zwischen zwei Artikeln (richtungsunspezifisch). Die Größe der Knoten ist proportional zum (Vernetzungs-)Grad des Knotens. Die Farbe der Knoten kodiert die Zugehörigkeit zu einem oder mehreren Teildiskursen (siehe Legende).

Auch eine bidirektionale Verbindung ist möglich. Graphen, in denen die Richtung einer Verbindung beachtet wird, heißen gerichtete Graphen. Wir betrachten Wörterbücher somit zunächst als ungerichtete Graphen.

Ein und derselbe Graph kann auf viele unterschiedliche Arten dargestellt werden. Dass in Abbildung 1 beispielsweise die Clique B, C, D, G unten links im Graphen steht, ist Zufall. Der Graph könnte genauso gut als Ganzes rotiert werden, bis die Clique oben rechts dargestellt wird. Es wird vielleicht auffallen, dass an

dieser Stelle der Vergleich zu Bahnhöfen nicht mehr trägt. Die Anordnung von Bahnhöfen kann selbstverständlich nicht einfach zufällig rotiert werden. Im Graphen nicht zufällig hingegen ist die Position des zentralen Knotens D. Dieser wird vom Zeichenalgorithmus immer zentral platziert, weil dies die optimale Position des Knotens in Relation zu allen anderen Knoten und den unter den Knoten bestehenden Verbindungen ist. Es würde keinen Sinn ergeben, wenn beispielsweise

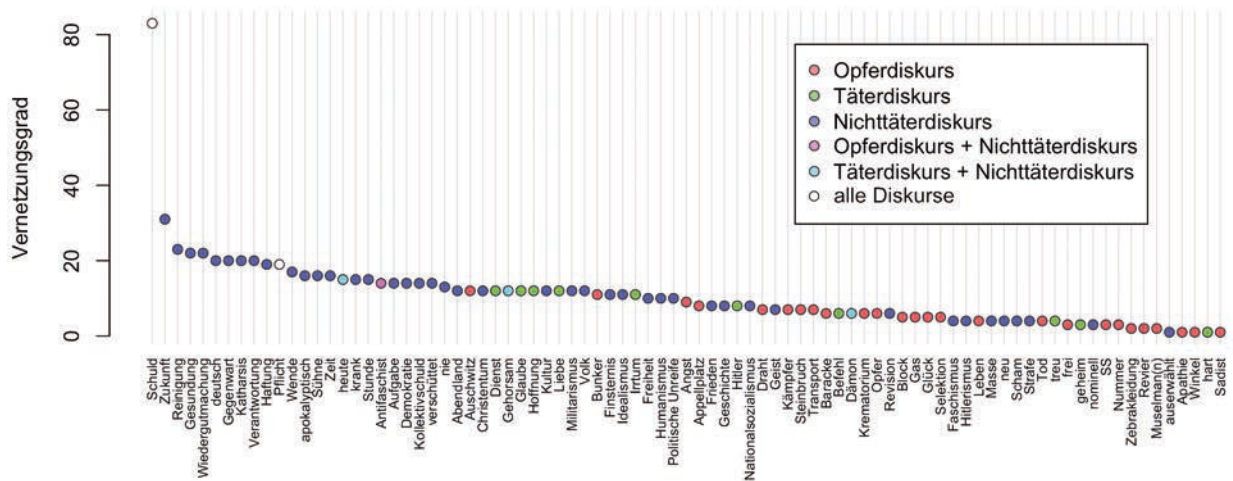


Abb. 3: (Vernetzungs-)Grade der Artikelknoten in der Übersicht. Die Farbe der Datenpunkte kodiert wiederum die Zugehörigkeit zu einem oder mehreren Teildiskursen (siehe Legende).

se der eher periphere Knoten F zentral platziert würde, weil sonst einige Kanten sehr lang würden und sich viele Kanten überschneiden müssten. Das heißt: Nicht nur die einzelnen Maßzahlen (z. B. der (Vernetzungs-) Grad eines Knotens) sind aussagekräftig, sondern auch die relative Platzierung eines Knotens, wie sie der Zeichenalgorithmus vornimmt, kann von Bedeutung sein. Im vorliegenden Artikel werden wir uns dies zunutze machen, wenn wir die Anordnung der Knoten im komplexen Wörterbuchgraphen interpretieren.

Das Wörterbuch als Graph: Anwendung und abgeleitete Erkenntnisse

Wir haben es im vorherigen Abschnitt bereits angedeutet. Die zentrale Frage dieses Artikels ist: Welche Schlüsse können gezogen werden, wenn wir ein Wörterbuch als ein Netz aus Artikeln und Verweisen, also als Graph, sehen? Datengrundlage ist die XML-Repräsentation des Diskurswörterbuchs zum Schuldiskurs 1945-1955. Wir erstellen also einen Wörterbuchgraphen, in dem jeder Artikel ein Knoten ist und jede Kante ein Verweis auf einen anderen Wörterbuchartikel. Abbildung 2 ist die grafische Repräsentation des Wörterbuchgraphens, die erwartungsgemäß nicht mehr so einfach ist wie der in Abbildung 1 gezeigte Graph unseres „Schienennetzes“.²

Der Wörterbuchgraph hat nicht nur deutlich mehr Knoten (85 Knoten) und Kanten (450 Kanten), sondern ist auch mit zusätzlicher Information angereichert:

- An jedem Knoten steht das Stichwort, das im entsprechenden Artikel beschrieben wird.

- Jeder Knoten hat eine bestimmte Farbe. Diese ergibt sich aus der Zuschreibung des jeweiligen Wortartikels zu einem oder mehreren Teildiskursen. Die Legende schlüsselt diese Farbzuordnungen auf. Weiße Knoten (*Pflicht* und *Schuld*) sind allen Teildiskursen zugeordnet.
- Die Knoten variieren in der Größe. Die Größe der Knoten ist proportional zum (Vernetzungs-)Grad des Knotens. Je mehr adjazente Kanten ein Knoten aufweist, d. h., je mehr Artikel auf den Knoten verweisen, desto größer ist er. Auch die Anzahl der ausgehenden Verweise wird hier gezählt.

Zunächst: Wer sehr gut mit dem Diskurswörterbuch zum Schuldiskurs vertraut ist, wurde evtl. schon hellhörig. Das Wörterbuch zum Schuldiskurs enthält 86 Artikel, der Graph aber nur 85 Knoten. Wenn jeder Artikel durch einen Knoten repräsentiert ist, sollte diese Anzahl eigentlich nicht voneinander abweichen. Vergleichen wir die Menge aller Artikel³ und die Namen der Knoten, tritt der fehlende Artikel zu Tage: Es ist der Knoten zum Artikel *Haltung*. Rufen wir die zugehörige Artikelseite⁴ in OWID auf, zeigt sich schnell, warum der Artikel keinen Knoten im Wörterbuchgraphen hat: Im Artikel *Haltung* wird auf keinen anderen Eintrag verwiesen und kein anderer Artikel verweist auf *Haltung*. Im Graphen wird somit kein Knoten für *Haltung* angezeigt.⁵ Bevor wir einzelne Muster im Wörterbuchgraphen auf qualitativer Ebene interpretieren, werden wir zunächst die (Vernetzungs-)Grade der einzelnen Artikelknoten zueinander in Beziehung setzen.

In Abbildung 3 sind alle Knoten hinsichtlich des Grades abgetragen. Es ist auf einen Blick zu sehen, dass der Eintrag *Schuld* am stärksten vernetzt ist. 83 Kanten liegen direkt am entsprechenden Knoten an. Diese Zahl setzt sich zusammen aus 56 ausgehenden Kanten

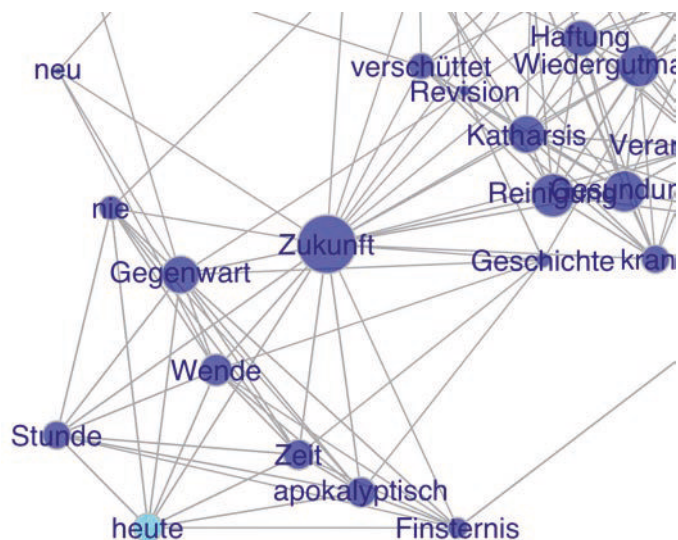


Abb. 4: Ausschnitt aus dem Gesamtwörterbuchgraphen rund um den Artikelknoten zum Eintrag *Zukunft*

(Verweise aus dem Artikel zu *Schuld*) und 27 eingehenden Kanten (Verweise zum Artikel *Schuld*). *Zukunft* ist der am zweitbesten vernetzte Eintrag. Hier liegen allerdings nur 31 Kanten an (17 ausgehende und 14 eingehende). Allein dieser Vergleich zwischen *Schuld* und *Zukunft* zeigt schon, wie zentral das Stichwort *Schuld* im Diskurs ist, den das Wörterbuch erfasst. Dies deckt sich auch mit der Graphendarstellung aus Abbildung 2, wo *Schuld* folgerichtig zentral angeordnet wird. Das Stichwort *Zukunft* bildet dort eine Art Nebenzentrum (in Abbildung 2 links unten im dargestellten Graphen), das aber nicht annähernd die Relevanz von *Schuld* erreicht und gut mit diesem eigentlichen Hauptzentrum verknüpft ist. Interessant ist auch der Vergleich des einzigen anderen Konzepts, das, außer *Schuld*, allen Teildiskursen zugeordnet ist, nämlich *Pflicht*. *Pflicht* nimmt Rang 11 bezüglich des Vernetzungsgrads ein (siehe Abbildung 3). Damit fällt es deutlich gegenüber dem zentralen Konzept *Schuld* ab, was die Vernetztheit mit dem Gesamtgraphen angeht. Im Vergleich zu der Darstellung des Gesamtwörterbuchs, die für die Online-Präsenz des Diskurswörterbuchs gewählt wurde⁶, zeigt sich hier ein Vorteil der Graphendarstellung. Die Darstellung für die Online-Präsenz basiert auf der Einteilung in Teildiskurse, die auch im vorliegenden Artikel bereits mehrfach angesprochen wurde. Da sowohl *Schuld* als auch *Pflicht* allen Teildiskursen zugeordnet sind, sind in dieser Darstellung beide Stichwörter zentral angeordnet. Die Analyse des Wörterbuchgraphen zeigt aber, dass *Pflicht* hinsichtlich der Relevanz für den Gesamtgraphen deutlich gegenüber *Schuld* abfällt. Wir wollen aber auch anmerken, dass die Darstellung, wie sie für die Online-Präsenz gewählt wurde, ebenfalls Vorteile gegenüber der Graphenrepräsentation hat: Sie ist vor allem deutlich einfacher und übersichtlicher. Auch die alphabetische Sortierung innerhalb der

Teildiskurse ließe sich in einem Graphen nur sehr schwer umsetzen, weil die Platzierung von anderen Prinzipien abhängt.

TEILDISKURSE BILDEN DISTINKTE GEBIETE IM GRAPHEN

Eine weitere Beobachtung ist aus Abbildung 2 abzuleiten: Die verschiedenen Teildiskurse (Opfer-, Täter- und Nichttäterdiskurse), denen jeder Artikel zugeordnet ist, werden vom Zeichenalgorithmus auch in distinkten „Gebieten“ des Graphen verortet: Blaue Knoten stehen nahe zusammen, grüne und rote ebenfalls. Dies geschieht deshalb, weil Artikel, die zum gleichen Teildiskurs gehören, untereinander weitaus stärker vernetzt sind als mit Artikeln, die anderen Teildiskursen zugeordnet sind. Folgerichtig platziert sie der Zeichenalgorithmus nahe beieinander, weil das die effektivste Darstellungsform ist. Das ist ein starkes Indiz dafür, dass sich die Einteilung der Artikel in verschiedene Teildiskurse konsistent in der Vernetzungsstruktur der Artikel untereinander niederschlägt.

DER GEMEINSAME BEZUG AUF *SCHULD* VERBINDET DIE TEILDISKURSBÜNDEL UNTEREINANDER

Alle diese Teildiskurs-„Bündel“ gruppieren sich um das zentrale Konzept *Schuld*, das als eine vermittelnde Instanz gesehen werden kann: Erst durch den gemein-

samen Bezug auf das Konzept *Schuld* ist eine Verbindung der verschiedenen Teildiskurse möglich. Zwar sind vereinzelt Verbindungen zwischen Artikeln distinkter Teildiskurse vorhanden⁷ (bspw. ein Verweis von *Auschwitz*, Opferdiskurs, zu *Nationalsozialismus*, Nichttäterdiskurs), doch das ist eben nicht die Regel. Die stärkere Vernetzung von Konzepten innerhalb ihrer jeweiligen Teildiskurse schlägt sich auch auf die in Abbildung 3 abgetragene Rangfolge nieder. Da es insgesamt mehr Konzepte gibt, die dem Teildiskurs der Nichttäter zugeordnet sind und dadurch mehr Verknüpfungen im Vergleich zu Teildiskursen mit weniger Mitgliedern entstehen, nehmen die Konzepte dieses Teildiskurses (blaue Datenpunkte) eher die oberen Ränge bezüglich des (Vernetzungs-)Grades ein.

Für eine weitere Interpretation wollen wir den in Abbildung 4 gezeigten Ausschnitt des Wörterbuchgraphen gesondert betrachten. Bei diesem Ausschnitt handelt es sich vornehmlich um Knoten aus dem Teildiskurs der Nichttäter rund um den Artikel zu *Zukunft* (in Abbildung 2 ist dieser Ausschnitt unten links verortet). Weiter oben haben wir bereits gezeigt, dass *Zukunft* der am zweitbesten vernetzte Knoten im Wörterbuchgraphen ist.

ZUKUNFT ALS VERMITTLERINSTANZ IN DEN RESTDISKURS

Abbildung 4 zeigt, warum dies der Fall ist: Um *Zukunft* herum gruppieren sich andere Zeitangaben (*Gegenwart*, *Zeit*, *Stunde*, *heute*) oder ähnliche Einträge, die zeitlich interpretiert werden können (*nie* und *Wende*),

aber auch *Finsternis* und *apokalyptisch* – als auf die Nachkriegsgegenwart referierende Kodierungen. All diese Artikel sind untereinander sehr gut vernetzt, wie die hohe Kantendichte in diesem Ausschnitt nahelegt. Allerdings weisen all diese Artikel keine oder nur wenige Verbindungen in den Restgraphen auf. Dafür sind sie aber alle mit dem Nebenzentrum *Zukunft* verbunden. *Zukunft* wiederum ist auch mit vielen anderen Artikeln gut vernetzt. Das Konzept *Zukunft* kann also als eine Art Vermittlerinstanz für die unten links davon abgetragenen Artikel in den Restgraphen verstanden werden. Diese Graphendarstellung entspricht dem Befund der Diskursanalyse: Der Diskurs der Nichttäter ist in hohem Maß zukunftsorientiert: „Das Lebensgefühl der geistigen Elite ist beherrscht von der Deutung der Gegenwart im Sinn einer, jede Möglichkeit der Gestaltung offenbarenden, ja, zu dieser Gestaltung auf-fordernden Zeit.“ (Kämper 2005, S. 178)

Schluss & Ausblick

Sehen wir Wörterbücher nicht als bloße Aneinanderreihung von Wortartikeln, sondern als ein komplexes Netz aus Verknüpfungen zwischen Konzepten, kann man zu neuen Einsichten über die Gesamtstruktur des beschriebenen Wortschatzausschnittes oder, wie in diesem Fall, die Zusammenhänge innerhalb eines Diskurses mit mehreren Teildiskursen kommen.

GRAPHENBASIERTE HERANGEHENSWEISE FÜHRT ZU NEUEN BEOBACHTUNGEN UND ERKENNTNISSEN

Für das Diskurswörterbuch zum Schuldiskurs 1945-1955 haben wir einige Aspekte herausgearbeitet, die anhand der graphentheoretischen Betrachtungsweise gelernt werden können. Das Konzept *Schuld* ist in der Tat auf mehreren Ebenen das zentrale Konzept. Stichwörter wie *Zukunft* oder *Pflicht* sind unter bestimmten Blickwinkeln ebenfalls relevant, kommen aber nicht an die Relevanz von *Schuld* heran.

Außerdem spiegelt sich die Einteilung in verschiedene Teildiskurse deutlich in der Vernetzungsstruktur und damit in der Platzierung in Bezug auf den Gesamtgraphen wider. Artikel zu Wörtern innerhalb eines bestimmten Teildiskurses verweisen entweder auf andere Artikel innerhalb des gleichen Teildiskurses oder eben auf das zentrale Konzept *Schuld*. Innerhalb des Teildiskurses der Nichttäter scheint das Stichwort *Zukunft* eine Art Vermittlerfunktion zu übernehmen. *Zukunft* integriert andere Stichwörter mit Bezug zu Zeit in den Restgraphen, sowohl im eigenen Teildiskurs als auch im Wörterbuchgraphen als ganzem.

Wir haben im vorliegenden Artikel nur einen kleinen Teil der Sachfragen erörtert, die man mit einer graphenbasierten Darstellung von Wörterbüchern bearbeiten kann. Potenzielle weitere Fragen im Wörterbuch zum Schulddiskurs 1945-1955 sind: Wie integrieren sich Konzepte, die mehreren Teildiskursen zugeordnet sind (wie bspw. *Antifaschist* oder *heute*) in den Restgraphen? Einige davon scheinen hauptsächlich an Stichwörter aus einem Teildiskurs „anzudocken“. *Heute* (unten links in Abbildung 2) bspw. reiht sich offenbar eher im Nichttäterdiskurs ein und hat keinerlei Verbindungen zu Konzepten des Täterdiskurses, dem es aber ebenfalls zugeordnet wurde. Eine weitere Frage wäre, wie relevant die Richtungen der Verbindungen sind. Könnte man aus diesen ebenfalls inhaltlich relevante Schlussfolgerungen ziehen? Im vorliegenden Beitrag wurden die Richtungen der Verknüpfungen nicht beachtet.

Die hier gezeigte Methode ließe sich ohne Weiteres auf das Wörterbuch zum Protestdiskurs 1967/68 übertragen, wir wollten uns hier aber auf eines der Diskurswörterbücher konzentrieren. Überträgt man die graphentheoretische Betrachtungsweise auf andere Arten von Wörterbüchern (wie in Müller-Spitzer / Wolfer 2015 geschehen), lassen sich noch andere Aussagen treffen. So lässt sich zum Beispiel zeigen, dass Stichwörter, die im kollaborativen Wörterbuchprojekt Wiktionary⁸ stärker vernetzt sind, auch häufiger in geschriebener Sprache vorkommen und im Wiktionary häufiger nachgeschlagen werden.

GRAPHENTHEORIE ERSETZT NICHT KLEIN- TEILIGE, KONZENTRIERTE ARBEIT AM WÖRTERBUCH

Zuletzt ein Hinweis bezüglich der Reichweite der hier vorgestellten Untersuchungen: Graphenbasierte Analysen von Wörterbüchern sind kein Mittel mit Alleinvertretungsanspruch. Gerade für Wörterbücher, in die ja bekanntlich viel Zeit für Recherche und Erstellung der Artikel fließt, gilt, dass eine graphenbasierte Betrachtungsweise niemals die konzentrierte, kleinteilige Arbeit am und mit dem sprachlichen Material ersetzen kann. Tatsächlich kann erst dann ein Graph erstellt werden, wenn Wortartikel zuvor sorgfältig manuell verknüpft wurden. Der hier vorgestellte Ansatz kann aber dazu dienen, eine andere Perspektive auf Wörterbücher und damit auf (Ausschnitte von) Sprache zu gewinnen.

KOMBINATION VON QUALITATIVER INTERPRETATION UND QUANTITATIVER ANALYSE IST GEWINNBRINGEND

Im vorliegenden Artikel haben wir versucht, qualitative Interpretation mit quantitativen Maßzahlen wie dem Vernetzungsgrad in Verbindung zu bringen. Von dieser Warte aus kann unser Beitrag auch als ein Integrationsversuch von qualitativer und quantitativer Forschungsarbeit gelesen werden. Gerade für die Sprachwissenschaft scheint ein solcher kombinierter Ansatz besonders gewinnbringend zu sein.

Anmerkungen

- ¹ Siehe hierzu den Eintrag in der Online-Version des Oxford Dictionary: <www.oxforddictionaries.com/definition/english/kick-the-bucket> [zuletzt aufgerufen am 12.06.2015]
- ² Unter <www1.ids-mannheim.de/fileadmin/lexik/bilder/schuld2_hd.png> (letzter Zugriff am 12.6.2015) ist eine hochauflösende Version des Graphen abrufbar. Alle Graphenberechnungen und -zeichnungen wurden mit der freien Statistikumgebung R (R Core Team, 2015) und dem Paket igraph (Csardi & Nepusz, 2006) erstellt.
- ³ Auf der Startseite des Online-Wortschatzinformationssystems Deutsch (OWID) <www.owid.de/docs/disk45/stw_idx.jsp> zum Diskurswörterbuch Schulddiskurs 1945-1955 ist eine Übersicht über alle beschriebenen Stichwörter abrufbar (letzter Zugriff am 12.06.2015).
- ⁴ <www.owid.de/artikel/309278> (letzter Zugriff am 12.06.2015).
- ⁵ Alternativ könnte man einen Knoten für *Haltung* am Rande des Wörterbuchgraphen darstellen – dann aber ohne verbindende Kanten in den Restgraphen.
- ⁶ Diese Darstellung ist unter <www.owid.de/docs/disk45/stw_idx.jsp> abzurufen (letzter Zugriff am 12.06.2015).
- ⁷ Solcherlei Verbindungen lassen sich am besten „aufspüren“, wenn man im Wörterbuchgraphen (Abbildung 2) nach Kanten sucht, die aus einem Teildiskurs-Bündel heraus und nicht zur Mitte streben, die also nicht in Richtung des zentralen Konzepts *Schuld* zeigen.
- ⁸ Die Startseite der deutschen Edition von Wiktionary befindet sich unter <<https://de.wiktionary.org/>> (letzter Zugriff am 11.6.2015).

Literaturverzeichnis

- Csardi, G. / Nepusz, T. (2006): The igraph software package for complex network research, InterJournal, Complex Systems 1695, <<http://igraph.org>> (letzter Zugriff am 17.6.2015).
- Kämper, Heidrun (2005): Der Schulddiskurs in der frühen Nachkriegszeit. Ein Beitrag zur Geschichte des sprachlichen Umbruchs nach 1945. Berlin, New York: de Gruyter.
- Kämper, Heidrun (2007): Opfer – Täter – Nichttäter. Ein Wörterbuch zum Schulddiskurs. Berlin, New York: de Gruyter. (Online: <www.owid.de/wb/disk45/einleitung.html>, letzter Zugriff am 17.6.2015)
- Kämper, Heidrun (2015): Diskurslexikografie als gesellschaftsbezogene Wortforschung. Vorstellung eines Wörterbuchkonzepts. In: Kilian, Jörg / Eckhoff, Jan (Hg.): Deutscher Wortschatz – beschreiben, lernen, lehren. Beiträge zur Wortschatzarbeit in Wissenschaft, Sprachunterricht, Gesellschaft. (Germanistik, Didaktik, Unterricht. Herausgegeben von Ina Karg. Band 13). Frankfurt/M., S. 21-38.
- Müller-Spitzer, Carolin / Wolfer, Sascha (2015): Vernetzungsstrukturen digitaler Wörterbücher. Neue Ansätze zur Analyse. Erscheint in: Lexicographica 2015, S. 173-199.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <www.R-project.org/> (letzter Zugriff am 17.6.2015).
- Turau, Volker (2009). Algorithmische Graphentheorie (3. Aufl.). München: Oldenbourg. ■